

Computational estimation of microRNA targets and binding site frequencies

Dannys J. Martínez-Herrera¹, Daniel Tabas-Madrid¹, Carlos Oscar Sánchez-Sorzano² and Alberto Pascual-Montano¹

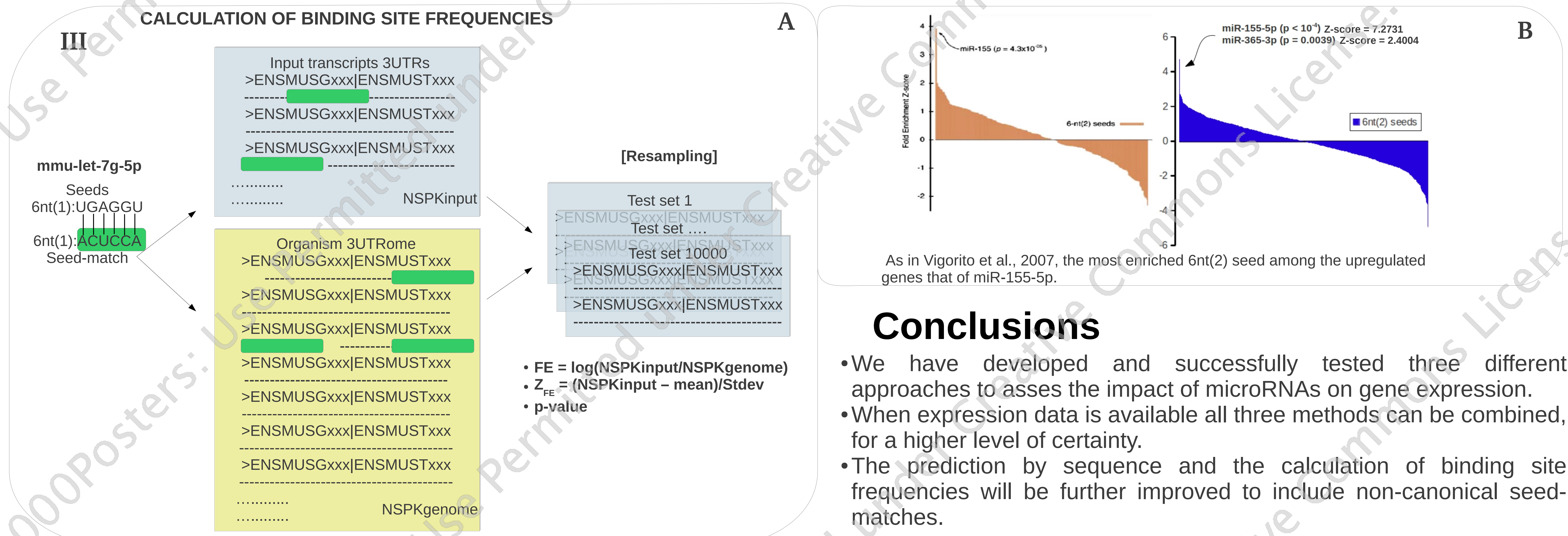
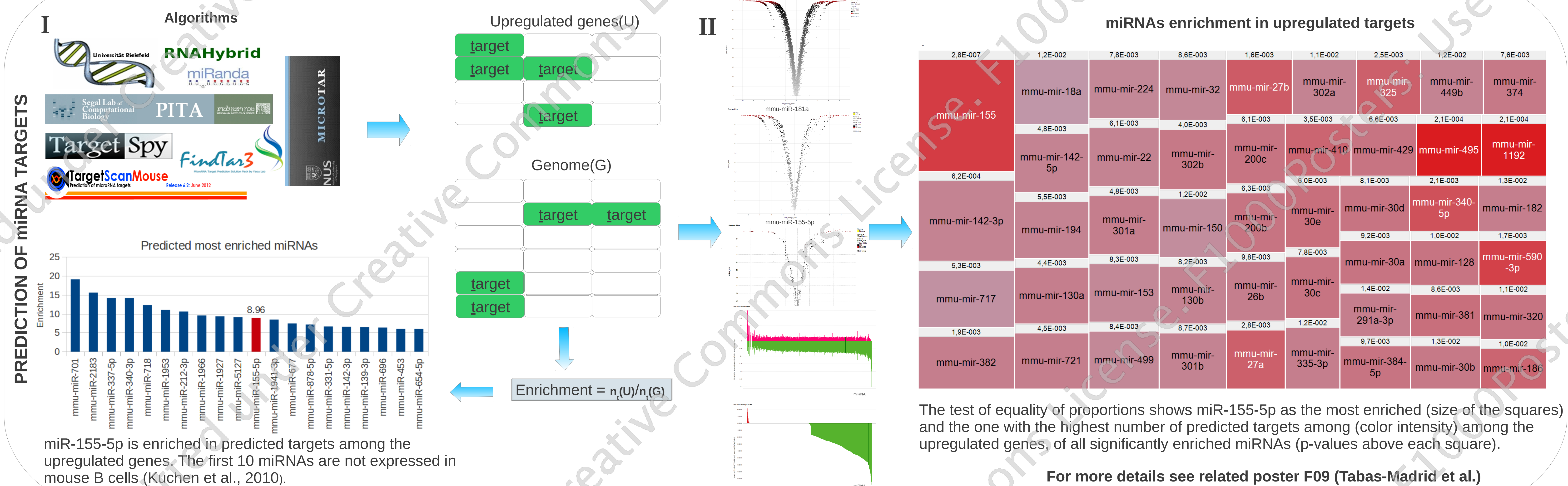
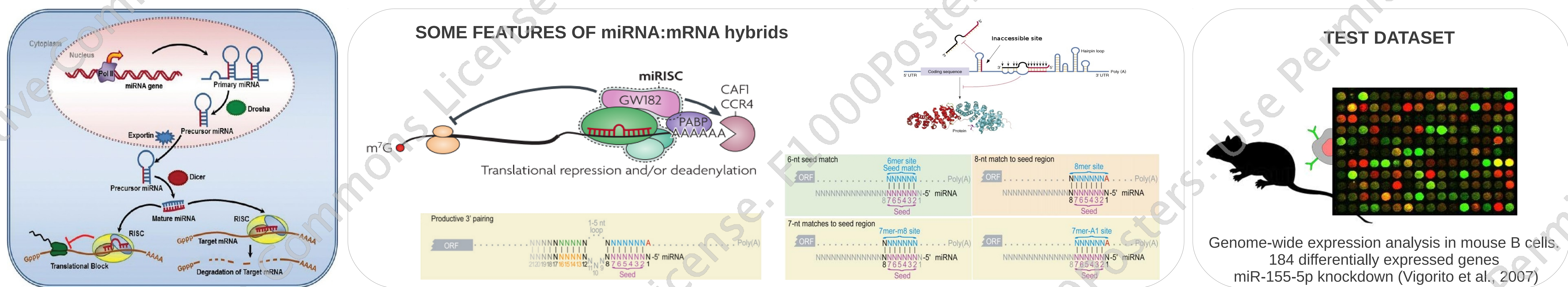
¹Functional Bioinformatics Group. National Center for Biotechnology-CSIC. Madrid, Spain

²BioComputing Unit. National Center for Biotechnology-CSIC. Madrid, Spain

e-mail: djmartinez@cnb.csic.es

Summary

MicroRNAs are ~22-nucleotide long molecules that regulate the expression of genes post-transcriptionally, first found in *C. elegans* by Lee and coworkers in 1993. However, there is still much to know about the factors that promote a successful miRNA-mRNA interaction. Here we propose three different methods to improve the predictions of miRNA targets. In a first approach, we perform the predictions of putative targets starting only from the sequences of miRNAs and the 3'UTRs of transcripts. For this purpose, we use seven algorithms based on different criteria, whose output is then filtered according to relevant biological data regarding the mode of action of microRNAs. When experimental transcriptomics data is available, we use it to determine the proportions of downregulated genes in the whole genome and compare them to the downregulated proportions of miRNA targets to discover miRNAs with significantly different ratios. Finally, we determine if downregulated mRNAs are enriched in putative miRNAs target sites by calculating the average number of sites per 1kb of 3' UTR sequences.



Conclusions

- We have developed and successfully tested three different approaches to assess the impact of microRNAs on gene expression.
- When expression data is available all three methods can be combined, for a higher level of certainty.
- The prediction by sequence and the calculation of binding site frequencies will be further improved to include non-canonical seed-matches.

DJ.M.H. acknowledges Fundación "la Caixa" for the economical support